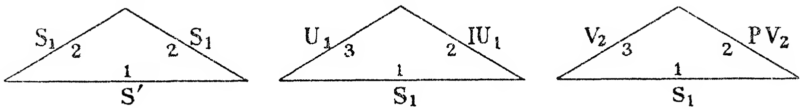
The point O which determines the dissection of the base is conveniently taken either at the middle point of the hypothenuse or at the centre of gravity. The exterior boundary which determines a second repeat is shown in dotted lines.

26. The isosceles triangle with vertical angle 120°, T. $\frac{2}{3}\pi$ can be assembled either on the contact system 11, 22, 33, or on 11, 23.

The classification is

$$\text{T. } \tfrac{2}{3}\pi. \quad \text{(i) } 1, 2, 2; \quad 1, 2, 3.$$
$$\text{T. } \tfrac{2}{3}\pi. \quad \text{(ii) } 1, 2, 3.$$

There are three symmetrical repeats:—



The point O may be suitably placed either at the vertex or at the mid point of the long side.

Correlation between Arrays in a Table of Correlations.

By C. Spearman.

(Communicated by Prof. L. N. G. Filon, F.R.S. Received December 28, 1921.)

Two recent papers in these Proceedings have dealt with certain problems of probability which are of very great importance for psychology in particular, but in themselves are quite general.* As both papers made reference to some of my own, the following considerations may not only be a further positive contribution to the topic, but incidentally serve to clear up some misunderstandings.

Suppose that the product moment coefficients of correlation have been determined between any set of variables $a, b, c, d, \ldots z$, and have been set forth in a square Table thus:—

	a	b	c	.	.	z
a		r_{ab}	r_{ac}	.	.	r_{az}
b	r_{ab}		r_{bc}	.	.	r_{bz}
c	r_{ac}	r_{bc}		.	.	r_{cz}
.	.	.	.	.	.	.
.	.	.	.	.	.	.
z	r_{az}	r_{bz}	r_{cz}	.	.	

* Maxwell Garnett, 1919; G. Thomson, 1919.

Both the said papers dealt with the problem as to how the correlation between parallel arrays of these coefficients (ignoring the coefficient that has no partner) are related to the inner constitution of the variable. One of the papers (Garnett) further inquired as to how this inner constitution is related to the general satisfaction of the equation,

$$r_{xy} = \lambda_{xz} r_{yz}, \text{ where } \lambda_{xz} \text{ is constant whilst } y \text{ takes all values except } x \text{ or } z.* \tag{1}$$

The present paper will endeavour to provide a general method for solving many of the cases at issue, and also will attempt actually to solve certain cases of primary importance.

1. Following Bravais (and also Pearson), we will represent each of the variables as a function of a number of elements uncorrelated with one another.

$$\left.\begin{aligned} a &= \phi\,(\epsilon_1 \epsilon_2 \ldots \epsilon_\kappa \epsilon_{\kappa+1} \ldots \epsilon_\nu) \\ b &= \psi\,(\epsilon_1 \epsilon_2 \ldots \epsilon_\kappa \epsilon_{\nu+1} \ldots \epsilon_\mu) \end{aligned}\right\} \tag{2}$$

where $\epsilon_1 \ldots \epsilon_\kappa$ are the elements common to both functions, whilst the rest occur only in one.

If, then, a and b are taken to have been measured respectively from their means as origins, and if—again with both Bravais and Pearson—we assume that these two functions are represented with sufficient approximation by Taylor's expansion to the first differentials, we get

$$a = A_1\epsilon_1 + A_2\epsilon_2 + \ldots A_\nu\epsilon_\nu, \text{ and similarly for } b, c, \ldots z. \tag{3}$$

For convenience, this may be written as $a = \sum_{\tau=1}^{\tau=\omega} (A_\tau \epsilon_\tau)$, where the Σ indicates summation over all the elements entering into *any* of the variables $a, b, \ldots z$, the coefficient A being taken as zero whenever the element is not present really.

Then, in the case of any particular observation of the variable a, say the mth, we get

$$a_m = \sum_{\tau=1}^{\tau=\omega} (A_\tau, \epsilon_{m\tau}), \tag{4}$$

and, letting S indicate summation with respect to the different observations of the same variable, we may for further convenience

denote $$A_\tau^2 \,.\, \mathop{S}_{m=1}^{m=n} (\epsilon_{m\tau}^2) \text{ by } A_\tau'^2. \tag{5}$$

* Otherwise expressed as $r_{px}/r_{qx} = r_{py}/r_{qy}$, for the psychological significance of which, see Spearman, 'Psychological Review,' vol. 27, p. 159.

Hereupon, bearing in mind that by hypothesis all products vanish,

$$\left.\begin{aligned} \mathop{\mathrm{S}}_{m=1}^{m=n} (a_m^2) &= \sum_{\tau=1}^{\tau=\omega} [\mathrm{A}_\tau^2 \,.\, \mathop{\mathrm{S}}_{m=1}^{m=n} (\epsilon_{m\tau}^2)] = \sum_{\tau=1}^{\tau=\omega} [\mathrm{A}_\tau'^2], \text{ and similarly} \\ \mathop{\mathrm{S}}_{m=1}^{m=n} (b_m^2) &= \sum_{\tau=1}^{\tau=\omega} [\mathrm{B}_\tau'^2] \\ \mathop{\mathrm{S}}_{m=1}^{m=n} (a_m b_m) &= \sum_{\tau=1}^{\tau=\omega} [\mathrm{A}_\tau' \mathrm{B}_\tau'] \end{aligned}\right\} \tag{6}$$

Writing next $\mathrm{A}' \Big/ \sqrt{\left(\sum_{\tau=1}^{\tau=\omega} [\mathrm{A}_\tau'^2]\right)}$ as A'', and similarly B'', (7)

we get
$$r_{ab} = \sum_{\tau=1}^{\tau=\omega} [\mathrm{A}''\mathrm{B}'']. \tag{8}$$

In similar fashion, the correlations of a and b with the rest of the variables $c \ldots z$ can be arranged in two columns as follows (dropping, for brevity, the indices to the symbols of summation)

$$\left.\begin{array}{ll} r_{ac} = \Sigma\,[\mathrm{A}''\mathrm{C}''], & r_{bc} = \Sigma\,[\mathrm{B}''\mathrm{C}''] \\ r_{ad} = \Sigma\,[\mathrm{A}''\mathrm{D}''], & r_{bd} = \Sigma\,[\mathrm{B}''\mathrm{D}''] \\ \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ r_{az} = \Sigma\,[\mathrm{A}''\mathrm{Z}''], & r_{bz} = \Sigma\,[\mathrm{B}''\mathrm{Z}''] \end{array}\right\} \tag{9}$$

If, next, R_{ab} denotes the coefficient of correlation between the two columns and x is any of the variable $c, d \ldots z$, then

$$\mathrm{R}_{ab} = r_{\Sigma[\mathrm{A}''\mathrm{X}''] \,.\, \Sigma[\mathrm{B}''\mathrm{X}'']} = r_{\Sigma[\mathrm{A}'\mathrm{X}''] \,.\, \Sigma[\mathrm{B}'\mathrm{X}'']},$$

since the factors $\sqrt{(\Sigma\,[\mathrm{A}'^2])}$ and $\sqrt{(\Sigma\,[\mathrm{B}'^2])}$ are constant and positive.

Hence, expanding by the formula for correlations of sums,* and using again Σ to denote summation in respect of all the elements from ϵ_1 to $\epsilon_{\omega 1}$, whilst σ and n refer to χ taking successively the values C, D ... Z,

$$\mathrm{R}_{ab} = \frac{\sum\limits_{\beta}[\mathrm{A}_\beta'\mathrm{B}_\beta'\sigma_{\mathrm{x}_\beta''}^2] + \sum\limits_{\beta\neq\gamma}[\mathrm{A}_\beta'\mathrm{B}_\gamma'\sigma_{\mathrm{x}_\beta''}\sigma_{\mathrm{x}_\gamma''}r_{\mathrm{x}_\beta''\mathrm{x}_\gamma''}]}{\sqrt{\left(\sum\limits_{\beta}[\mathrm{A}'^2\sigma_{\mathrm{x}_\beta''}] + 2\sum\limits_{\beta\neq\gamma}[\mathrm{A}_\beta'\mathrm{A}_\gamma'\sigma_{\mathrm{x}_\beta''}\sigma_{\mathrm{x}_\gamma''}r_{\mathrm{x}_\beta''\mathrm{x}_\gamma''}]\right)} \sqrt{\left(\sum\limits_{\beta}[\mathrm{B}_\beta'^2\sigma_{\mathrm{x}_\beta''}^2] + 2\sum\limits_{\beta\neq\gamma}[\mathrm{B}_\beta'\mathrm{B}_\gamma'\sigma_{\mathrm{x}_\beta''}\sigma_{\mathrm{x}_\gamma''}r_{\mathrm{x}_\beta''\mathrm{x}_\gamma''}]\right)}}. \tag{10}$$

2. Now, this last equation (10) has numerous applications and simplifications according to any special conditions that may be introduced. We will here apply it to the most important condition of all, that is, where all the correlations derive from one and the same element alone, *e.g.*,

$$\begin{aligned} a &= \mathrm{A}_1\epsilon_1 + \mathrm{A}_\nu\epsilon_\nu + \ldots \\ b &= \mathrm{B}_1\epsilon_1 + \mathrm{B}_\mu\epsilon_\mu + \ldots \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \\ z &= \mathrm{Z}_1\epsilon_l + \mathrm{Z}_\theta\epsilon + \ldots \end{aligned}$$

ϵ_ν, ϵ_μ, and ϵ_θ, etc., being always different.

* Spearman, 'Brit. J. Psych.,' vol. 5, p. 419 (1913).

Then, no summation with respect to τ is required for (8) and (9). For (10), the second term in the numerator, as well as under each of the roots in the denominator, vanishes. We have, therefore,

$$R_{ab} = \frac{A'B'\sigma_{x''}^2}{\sqrt{(A'^2\sigma_{x''}^2)}\sqrt{(B'^2\sigma_{x''}^2)}} = \pm 1, \tag{11}$$

according as A′ and B′ are of the same or different signs. This is in agreement with the conclusions reached previously (by way of Udny Yule's formula for partial coefficients), that the present assumption of there being only one single common element leads to equation (1).*

In this connection, consider also the case where the variables $a, b, \ldots z$ have several common elements, but always the same and in similar proportions, *e.g.*,

$$p = \lambda_{pq} Q_\zeta \epsilon_\zeta + \ldots + \lambda_{pq} Q_\theta \epsilon_\theta + \ldots P_\rho \epsilon_\rho + \ldots + P_\xi \epsilon_\xi, \text{ where } \epsilon_\zeta \ldots \epsilon_\theta$$

are common to all the variables, whilst no others are common to any: λ_{pq} is constant for different elements. Then, since X_β is always in the same proportion to X_γ, so also is X_β'' to X_γ'' whereby $r_{x_\beta'' x_\gamma''} = 1$. Hence using (10)

$$R_{ab} = \frac{\Sigma[A'B'\sigma_{x''}^2] + \sum_{\beta\neq\gamma}[A_\beta' B_\gamma' \sigma_{x_\beta''}\sigma_{x_\gamma''}]}{\sqrt{(\Sigma[A'^2\sigma_{x''}^2] + \sum_{\beta\neq\gamma}[A_\beta' A_\gamma' \sigma_{x_\beta''}\sigma_{x_\gamma''}])}\sqrt{(\Sigma[B'^2\sigma_{x''}^2] + \sum_{\beta\neq\gamma}[B_\beta' B_\gamma' \sigma_{x_\beta''}\sigma_{x_\gamma''}])}}$$

$$= \frac{\Sigma[A'\sigma_{x''}] \,.\, \Sigma[B'\sigma_{x''}]}{\sqrt{(\{\Sigma[A'\sigma_{x''}]\}^2)}\sqrt{(\{\Sigma[B'\sigma_{x''}]\}^2)}} = \pm 1, \text{ as before.} \tag{12}$$

3. So far, our inferences have been from the inward structure of the variables to the values in the correlational table. But even more important is the inverse direction of inference, especially when applied to the same condition as before. That is to say, having seen that the fact of only one independent element being common to two or more variables proves both $R_{xy} = 1$ and also equation (1), the question now arises as to whether either of these equations proves that there can be only one element.

As regards $R_{xy} = 1$, this question must be reserved for another occasion. But as regards equation (1) the answer can be shown here to be quite generally affirmative.

The question is tantamount to asking whether, on assuming (1), each of the variables can be reduced to the form

$$a = f_a \eta + \delta_a \tag{13}$$

where 1. f_a, f_b, etc., are constant for all particular values of a, b, etc.

2. η is an element common to all the variables.

3. δ_a, δ_b, etc., are uncorrelated with η,

4. δ_a, δ_b, etc., are uncorrelated with each other.

* Hart and Spearman, 'Brit. J. Psych.,' vol. 5, p. 58 (1912).

Now, we can always write any of the variables, say a, so as to satisfy 1 and 2, giving to f_a and to η any values we please, so long as, and only so long as, $\delta_a = a - f_a\eta$.

But in order to fulfil the third condition, we must also have

$$0 = r_{(a-f_a\eta)(f_a\eta)} \text{ which is equivalent to } 0 = \sigma_a r_{a_\eta} - f_a \sigma_\eta,^*$$

which $= r_{a\eta} - f_a$, if this time we choose the units so that $\sigma_a = \sigma_b = \ldots \sigma_\eta = 1$. Hence we can fulfil the third condition by, and only by, making $f_a = r_{a\eta}$.

In such manner, the first three conditions can be satisfied for any set of variables whatever. But there remains the fourth condition, which will be satisfied if, and only if, we also obtain

$0 = r_{\delta_a\delta_b} = r_{(a-f_a\eta)(b-f_b\eta)}$ and therefore

$= r_{ab} - f_a r_{b\eta} - f_b r_{a\eta} + f_a f_b,$† which, on replacing the f's by their values from (13)

$$= r_{ab} - r_{a\eta} r_{b\eta}. \tag{14}$$

This is effected easily in the case where the set of variables entering into the table of coefficients is very large. For then we need only choose for η the value of $\overset{x=z}{\underset{x=a}{\mathbf{S}}}(x)$, where the summation is over the said large number of variables, and there is a change of unit so as to make $\sigma_\eta = 1$. For now we get

$r_{a\eta} = \dfrac{\sqrt{(m)}\,\bar{r}_{ax'}}{\sqrt{(1+[m-1]\,\bar{r}_{x'x''})}},$† where $x' \neq x''$ denote any of the variables including a itself, whilst m is the number of these variables

$= \dfrac{\bar{r}_{ax'}}{\sqrt{(\bar{r}_{x'x''})}}$ approximately, since m is by assumption very large.

Hence,
$$r_{a\eta} r_{b\eta} = \frac{\bar{r}_{ax'}\bar{r}_{bx'}}{r_{x'x''}} \tag{15}$$

But by (1),
$$r_{ab} = \frac{r_{ax} r_{bx^*}}{r_{xx^*}} = \frac{\overline{r_{ax} r_{bx^*}}}{\bar{r}_{x'x''}} = \frac{\bar{r}_{ax'}\bar{r}_{bz} + \sigma_{r_{ax}}\sigma_{r_{bx}} r_{r_{ax}r_{bx^*}}}{\bar{r}_{x'x''}}. \tag{16}$$

Further, the numerator of $r_{r_{ax}r_{bx^*}}$ can be written as $\overline{(r_{ax} - \bar{r}_{ax})(r_{bx^*} - \bar{r}_{bx})}$, which, owing to the largeness of m, approximates to the value it would have if x and x^* took all values quite independently of each other. And this is zero.

Again owing to the largeness of m, $\bar{r}_{ax'}$ and $\bar{r}_{bx'}$ approximate respectively to $\bar{r}_{ax}$ and $\bar{r}_{bx}$.

* Spearman, 'Brit. J. Psych.,' vol. 5, p. 419.

† See preceding footnote.

Accordingly, r_{ab} approximates to $\frac{\bar{r}_{ax'} \cdot \bar{r}_{bz'}}{\bar{r}_{x'x''}}$, which by (15) $= r_{a\eta} r_{b\eta}$, so that $r_{\delta_a \delta_b} = 0$ and condition 4 is fulfilled as required. (17)

This not only corroborates the result obtained in the above mentioned paper of Garnett, but dispenses with his particular and precarious assumption of "normal" frequency distributions, so that it extends the theorem to one of perfect generality, except that m must be very large.

In order to fulfil condition 4, freed even from *this* limitation, the constitution of η must be more complicated. Let us choose for it the value expressed in the following determinant

$$\begin{vmatrix} p\sqrt{(S/M_i i)} & p\mu_a a & p\mu_b b & . & p\mu_z z \\ -1 & \mu_a^2 - 1 & 0 & . & 0 \\ -1 & 0 & \mu_b^2 - 1 & . & 0 \\ . & . & . & . & . \\ -1 & 0 & 0 & . & \mu_z^2 - 1 \end{vmatrix}$$

where $\mu_a \equiv 1/\sqrt{(\lambda_{aq} r_{aq})}$, λ having a meaning as in (1), so that μ_a retains the same value whatever variable may be taken as q,

$i \equiv$ *any new* variable uncorrelated with all the others ($\sigma = 1$),*

$M_i \equiv$ the complementary minor of the first element† in the first row,

$S \equiv$ the sum of such minors for all elements in the first row,

and p is such a value as will make $\sigma_\eta = 1$.

This gives us, expanding the determinant according to the elements of the first row,‡

$$r_{a\eta} = \frac{\mu_a M_a + \mu_b M_b r_{ab} + \ldots + \mu_z M_z r_{az}}{\sqrt{(\mu_a^2 M_a^2 + \ldots + \mu_z^2 M_z^2 + M_i S + 2\mathbf{S}_x \mathbf{S}_y [\mu_x \mu_y M_x M_y r_{xy}])}}.$$

But by (1)

$$r_{xy} = \lambda_{xq} r_{yq} = \lambda_{yq} r_{xq} = \sqrt{(\lambda_{xq} r_{xq})} \sqrt{(\lambda_{yq} r_{yq})} = 1/\mu_x \; 1/\mu_y, \text{ so that}$$

$$r_{a\eta} = \frac{1}{\mu_a} \frac{(\mu_a^2 - 1) M_a + S'}{\sqrt{([\mu_a^2 - 1] M_a^2 + \ldots + [\mu_z^2 - 1] M_z^2 + M_i S + S'^2)}} \text{ where } S' = S - M_i.$$

But, on considering the determinant, clearly $(\mu_x^2 - 1) M_x = M_i$. Substituting M_i accordingly

$$r_{a\eta} = \frac{1}{\mu_a} \frac{S}{\sqrt{(M_i S' + S'^2 + M_i S)}} = \frac{1}{\mu_a}. \tag{18}$$

* This is here assumed to be possible, at any rate if the present theorem is only applied to a finite number of variables. The case of an infinite number has already been demonstrated by (17).

† Of course, "element" is here no longer used with the same meaning as previously in this paper.

‡ Spearman, 'Brit. J. Psych.,' vol. 5, p. 419.

Consequently $r_{ab} - r_{a\eta} r_{b\eta}$ or $r_{\delta_a \delta_b} = 0$. And extending this result to all the other variables, all of them become actually reduced to the required form of $f_x \eta + \delta_x$, fulfilling all the required conditions.

The practical bearings of some of the foregoing results may be illustrated by reference to the following Table of coefficients, obtained from mental tests applied to 757 children* :—

	Mathematical judgment.	Controlled association.	Literary interpretation.	Selective judgment.	Spelling.
Mathematical judgment ...	—	0·485	0·400	0·397	0·295
Controlled association	0·485	—	0·397	0·397	0·247
Literary interpretation ...	0·400	0·397	—	0·335	0·275
Selective judgment	0·397	0·397	0·335	—	0·195
Spelling	0·297	0·247	0·275	0·195	—

The median value of R_{ab} (even without any correction to compensate for the disturbance by sampling errors) works out at over 0·98. It thus accords very exactly with the value obtained in (11) and suggests the applicability of the theorem demonstrated in 3 whereby all the correlations will be exclusively due to some single common element.

* The experiments are by Bonser. They are quoted with reference in 'Brit. Journ. Psych.,' vol. 5, p. 62 (1912).